

Integration by differentiation: new proofs, methods and examples

Ding Jia (贾丁)^{1,2}, Eugene Tang³, Achim Kempf^{1,2}

¹Departments of Applied Mathematics
University of Waterloo, Ontario N2L 3G1, Canada

² Perimeter Institute for Theoretical Physics
Waterloo, ON, N2L 2Y5, Canada

³Division of Physics, Mathematics and Astronomy
California Institute of Technology, Pasadena, CA 91125, United States

E-mail: d7jia@uwaterloo.ca

Abstract. Recently, new methods were introduced which allow one to solve ordinary integrals by performing only derivatives. These studies were originally motivated by the difficulties of the quantum field theoretic path integral, and correspondingly, the results were derived by heuristic methods. Here, we give rigorous proofs for the methods to hold on fully specified function spaces. We then illustrate the efficacy of the new methods by applying them to the study of the surprising behavior of so-called Borwein integrals.

1. Introduction

Integration is generally harder to perform than differentiation. In particular, functional integrals, i.e., Feynman path integrals, are harder to work with, or even to define, than functional derivatives. See, for example, [1, 2, 3, 4, 5]. In this general context, methods have recently been found that allow one to integrate, or path integrate, by performing only derivatives, or functional derivatives, respectively. [6, 7]. While these methods were originally introduced to express quantum field theoretical path integrals in terms of much easier-to-handle functional derivatives, it was found that the new integration-by-differentiation techniques also add valuable tools to the toolbox for the evaluation of ordinary integrals.

So far, however, the domain of validity of the new methods has been little explored. Therefore, we here give rigorous proofs which show that the new methods hold for certain large classes of function spaces. Further, we illustrate the power of the new methods by applying them to sequences of so-called Borwein integrals. These integrals are nontrivial to evaluate and are known to exhibit rather curious behaviors. We demonstrate how the new methods allow one to evaluate even the normally complicated Borwein integrals quickly and transparently.

2. Integration by differentiation

Assume that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a function which has a power series expansion, $f(x) = \sum_{k=0}^{\infty} a_k x^k$, whose radius of convergence is nonzero. For any choice of $r \in \mathbb{C}$, we define a differential operator which we will denote by $f(r\partial_x)$. By definition, $f(r\partial_x)$ acts on smooth functions $\varphi(x)$ as

$$f(r\partial_x)\varphi(x) := \lim_{N \rightarrow \infty} \sum_{k=0}^N a_k (r\partial_x)^k \varphi(x), \quad (1)$$

whenever the latter limit is convergent. For convenience, we defined $\partial_x := d/dx$. For example, if $f(x) = \exp(x)$ then $f(r\partial_x) = \exp(r\partial_x)$, which can act on smooth functions such as $\varphi(x) = \sin(x)$, giving

$$f(r\partial_x)\varphi(x) = e^{r\partial_x} \sin(x) = \lim_{N \rightarrow \infty} \sum_{k=0}^N \frac{(r\partial_x)^k}{k!} \sin(x) = \sin(x + r). \quad (2)$$

The last step follows from the Taylor expansion.

Using this definition of $f(r\partial_x)$, the following integration by differentiation methods have been introduced [6, 7]:

$$\int_a^b f(x) dx = \lim_{y \rightarrow 0} f(-i\partial_y) \frac{e^{iby} - e^{iay}}{iy}, \quad (3)$$

$$\int_a^b f(x) dx = \lim_{y \rightarrow 0} f(\partial_y) \frac{e^{by} - e^{ay}}{y}, \quad (4)$$

$$\int_0^{\infty} f(x) dx = \lim_{y \rightarrow 0^+} f(-\partial_y) \frac{1}{y}, \quad (5)$$

$$\int_{-\infty}^0 f(x) dx = \lim_{y \rightarrow 0^+} f(\partial_y) \frac{1}{y}, \quad (6)$$

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{y \rightarrow 0^+} \left(f(\partial_y) + f(-\partial_y) \right) \frac{1}{y}, \quad (7)$$

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{y \rightarrow 0^+} 2\pi f(-i\partial_y) \delta(y). \quad (8)$$

Further, these equations are the “zero frequency” cases of the following general formulas for the Fourier and Laplace transforms:

$$\int_a^b f(x)e^{ixy} dx = f(-i\partial_y) \frac{e^{iby} - e^{iay}}{iy}, \quad (9)$$

$$\int_a^b f(x)e^{xy} dx = f(\partial_y) \frac{e^{by} - e^{ay}}{y}, \quad (10)$$

$$\int_0^\infty f(x)e^{-xy} dx = f(-\partial_y) \frac{1}{y}, \quad (11)$$

$$\int_{-\infty}^0 f(x)e^{xy} dx = f(\partial_y) \frac{1}{y}, \quad (12)$$

$$\int_{-\infty}^\infty f(x)e^{ixy} dx = 2\pi f(-i\partial_y)\delta(y). \quad (13)$$

For example, here is a calculation based on equation (4):

$$\int_a^b xe^{-x} dx = \lim_{y \rightarrow 0} e^{-\partial_y} \partial_y \frac{e^{by} - e^{ay}}{y} \quad (14)$$

$$= \lim_{y \rightarrow 0} e^{-\partial_y} \frac{(by - 1)e^{by} - (ay - 1)e^{ay}}{y^2} \quad (15)$$

$$= \lim_{y \rightarrow 0} \frac{(by - b - 1)e^{b(y-1)} - (ay - a - 1)e^{a(y-1)}}{(y - 1)^2} \quad (16)$$

$$= -(b + 1)e^{-b} + (a + 1)e^{-a}. \quad (17)$$

We used here the fact that $e^{a\partial_y} f(y) = f(y + a)$, i.e., that $e^{a\partial_y}$ acts as a translation operator (see Lemma 5.1). While the integration by differentiation formulas were originally derived heuristically, our aim here is to prove them rigorously and establish conditions under which they apply.

3. Main Results

In this section we present a list of propositions that put the above integration by differentiation methods on a rigorous footing. The proofs are given in Section 5. Throughout, we will continue to define $f(r\partial_x)$ through the power series expansion of f , unless otherwise mentioned.

Indeed, before turning to the main results, let us briefly discuss an alternative way in which we could define the function of a derivative, namely not through a power series expansion but instead through the spectral calculus. Concretely, using the spectral theorem, we could define

$$f(-i\partial_x) \equiv \mathcal{F} M_{f(-x)} \mathcal{F}^{-1}, \quad (18)$$

where \mathcal{F} is the Fourier transform, which we define through:

$$\mathcal{F}[f](y) = \hat{f}(y) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{ixy} dx, \quad (19)$$

Here, $M_{f(-x)}$ is the multiplication operator $M_{f(-x)}\varphi(x) = f(-x)\varphi(x)$. Note the unconventional definition of the Fourier transform here, with kernel e^{ixy} instead of e^{-ixy} . This is simply to minimize unnecessary factors of -1 later on.

The reason why we chose to define the function of a derivative through the power series definition in (1) is that, for practical calculations, the power series definition is generally more useful than the spectral definition. This is because the spectral definition requires a priori the very Fourier transform that the new methods are meant to find. Nevertheless, it is worth noting that the spectral definition of $f(r\partial_x)$ has the advantage of generally admitting a much larger class of functions for which the integration by differentiation formulas can hold. This is because the definition (18) does not need the function f to be analytic, or even continuous. For example, for studies of equations that hold for path integrals, such as perturbative expansions, that larger class of admissible functions might provide greater flexibility.

3.1. Finite intervals

We begin with the equations that are the most straightforward to establish rigorously, namely the formulas which involve only integrals over finite intervals. The basic results are summarized in the following proposition:

Proposition 3.1. *Suppose that $f(x)$ has a power series expansion $f(x) = \sum_{k=0}^{\infty} a_k x^k$ with a radius of convergence covering $[a, b]$. Then equations (3), (4), (9), and (10) hold.*

3.2. Laplace-type methods

From the previous finite integral equations, we may take the limits of integration to $\pm\infty$. As we shall show in section 5, careful taking of the limit yields rigorous versions of (11) and (12):

Proposition 3.2. *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is entire and Laplace transformable on \mathbb{R}_+ , then*

$$\int_0^{\infty} f(x) e^{-xy} dx = \lim_{a \rightarrow \infty} f(-\partial_y) \frac{1 - e^{-ay}}{y}, \quad (20)$$

for all $y \in \mathbb{R}_+$ for which the integral is convergent. Likewise, if f is entire and Laplace transformable[‡] on \mathbb{R}_- , then

$$\int_{-\infty}^0 f(x) e^{xy} dx = \lim_{a \rightarrow \infty} f(\partial_y) \frac{1 - e^{-ay}}{y} \quad (21)$$

for all $y \in \mathbb{R}_+$ for which the integral is convergent.

[‡] By Laplace transformable on \mathbb{R}_- , we mean that there exists some $y_0 \in \mathbb{R}_+$ such that $\int_{-\infty}^0 f(x) e^{xy} dt$ is convergent for all $y > y_0$.

If the function f is entire and integrable on the real or half-lines, then we may first take the zero frequency limit $y \rightarrow 0$ before taking the limits of integration to $\pm\infty$. This yields the following proposition:

Proposition 3.3. *If f is entire and integrable on the half-lines \mathbb{R}_+ or \mathbb{R}_- , then we have*

$$\int_0^\infty f(x) dx = \lim_{a \rightarrow \infty} \lim_{y \rightarrow 0} f(-\partial_y) \frac{1 - e^{-ay}}{y}, \quad (22)$$

$$\int_{-\infty}^0 f(x) dx = \lim_{a \rightarrow \infty} \lim_{y \rightarrow 0} f(\partial_y) \frac{1 - e^{-ay}}{y}, \quad (23)$$

respectively. Then, if $f : \mathbb{R} \rightarrow \mathbb{R}$ is entire and integrable on \mathbb{R} , we have

$$\int_{-\infty}^\infty f(x) dx = \lim_{a \rightarrow \infty} \lim_{y \rightarrow 0} [f(\partial_y) + f(-\partial_y)] \frac{1 - e^{-ay}}{y}, \quad (24)$$

$$\int_{-\infty}^\infty f(x) dx = \lim_{a \rightarrow \infty} \lim_{y \rightarrow 0} f(-i\partial_y) 2a \operatorname{sinc}(ay). \quad (25)$$

Let us now turn our attention to equations (11) and (12). Looking at Proposition 3.2, we notice a remarkable similarity between equations (11) - (12), and equations (20) - (21). Indeed, if we could exchange the limit $a \rightarrow \infty$ and the operator $f(-\partial_y)$ in equations (20) and (21), then we would immediately obtain (11) and (12). However, these operations in general cannot be exchanged, and so there is some subtlety in the sense in which equations (11) and (12) hold:

From the right-hand side of (11), we can see that the power series differential operator acting on $1/y$ gives us a Laurent series in y instead of the usual Taylor series. This is natural, considering that the convergence of the Laplace transform becomes better with increasing y . Given the fact that equation (11) holds with a Laurent series, it can happen that the domain of convergence for equation (11) is not the full domain of convergence of the corresponding Laplace transform.

For example, consider the application of equation (11) to the function $f(x) = e^{-x}$. This function is Laplace transformable, and the Laplace transform

$$\mathcal{L}[e^{-x}](y) = \int_0^\infty e^{-x} e^{-xy} dx = \frac{1}{y+1} \quad (26)$$

is convergent for all $y > -1$. If we evaluate the same example with equation (11), then we find

$$e^{\partial_y} \left(\frac{1}{y} \right) = \sum_{k=0}^\infty \frac{1}{k!} \frac{d^k}{dy^k} \left(\frac{1}{y} \right) = \sum_{k=0}^\infty (-1)^k \frac{1}{y^{k+1}}. \quad (27)$$

This is a proper Laurent series which equals $1/(y+1)$, as it should, but which is only convergent for $|y| > 1$. Thus we see that the two domains differ and that, therefore, these problems warrant further investigation.

The underlying issue here is the fact that, in equation (11), the operator $f(-\partial_y)$ is acting on $1/y$, which has a pole at $y = 0$. According to our new Proposition 3.2, the use of the equation (20), which is the regularized version of equation (11), must succeed in yielding the full solution.

Let us verify this: Indeed, in contrast to $1/y$, which is singular, the function $(1 - e^{-ay})/y$ is an entire function. Applying (20) to our example $f(x) = e^{-x}$ then gives us

$$\lim_{a \rightarrow \infty} e^{\partial_y} \left(\frac{1 - e^{-ay}}{y} \right) = \lim_{a \rightarrow \infty} \frac{1 - e^{-a(y+1)}}{y+1} = \frac{1}{y+1}, \quad (28)$$

where the final limit is convergent for $y > -1$, which means that we obtain the Laplace transform on its full domain. Note that because of the fact that $(1 - e^{-ay})/y$ is entire, we were able to use e^{∂_y} as the translation operator (see Lemma 5.1). We were not able to do this for $1/y$. Moreover, for each finite value of a , the resulting function

$$e^{\partial_y} \left(\frac{1 - e^{-ay}}{y} \right) = \frac{1 - e^{-a(y+1)}}{y+1} \quad (29)$$

is an entire function of y . Therefore, equation (20) avoids the convergence problem by obtaining its results as the limit of entire functions, bypassing the need for a Laurent series expansion. The two equations will of course agree within their common domain of convergence, and one may take the view of equation (20) as giving an analytic continuation of equation (11).

Now for practical purposes, equation (11) is easier to use than equation (20). The problem with (11), however, is that as we just saw, its domain of convergence is nontrivial. The following proposition provides a practical criterion with mild conditions on f for the use of equation (11):

Proposition 3.4. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be an entire function given by*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k. \quad (30)$$

Let us also define the function $\tilde{f}(x)$ by

$$\tilde{f}(x) = \sum_{k=0}^{\infty} |a_k| x^k. \quad (31)$$

Suppose that \tilde{f} is Laplace transformable, such that the integral

$$\mathcal{L}[\tilde{f}](y) = \int_0^{\infty} \tilde{f}(x) e^{-xy} dx \quad (32)$$

is convergent for all $y > y_0 \geq 0$. Then we have

$$\mathcal{L}[f](y) = \int_0^{\infty} f(x) e^{-xy} dx = f(-\partial_y) \frac{1}{y}, \quad (33)$$

which holds for all $y > y_0$. The statement for the Laplace transform on $(-\infty, 0]$ holds analogously.

Applying this proposition to $f(x) = e^{-x}$, we have $\tilde{f}(x) = e^x$, and so the Laplace transform for $\tilde{f}(x)$ is convergent for all $y > y_0 = 1$. Indeed, this is precisely the domain of convergence we found for the Laurent series before.

Finally, let us turn towards the zero frequency case of equations (11) and (12), given by equations (5) - (7). From our previous example, we see that the domain of convergence for $\tilde{f}(y)$ can easily be smaller than the positive half line, i.e., with $y_0 > 0$. In this case, $y = 0$ is not a limit point for the domain of convergence, and therefore the limit $y \rightarrow 0$ to obtain an integration formula cannot be taken. Therefore, to be safe, the two equations (5) - (7) must, in general, be used in their regularized forms given by equations (20) - (21) and (24).

Fortunately, as we will now show, for certain classes of functions, these regularizations are not needed and the integration by differentiation equations (5) - (7) produce the correct answers even when used directly. This is the content of the next proposition.

Proposition 3.5. *Suppose $f : \mathbb{R} \rightarrow \mathbb{R}$ is an entire and integrable function of the form*

$$f(x) = \sum_{j=1}^N c_j e^{-b_j x} x^{n_j}, \quad (34)$$

with $b_i > 0$, $c_j \in \mathbb{C}$, and $n_j \in \mathbb{Z}$. Then equations (5) - (7) hold with the action of $f(\pm \partial_y)$ on the right-hand side given formally (instead of as a power series). Namely, we define $e^{b\partial_y}$ to act by translation

$$e^{b\partial_y} \varphi(y) = \varphi(y + b),$$

and ∂_y^n to give the n th derivative ($n > 0$) or anti-derivative ($n < 0$).

Moreover, when anti-differentiating the choice of representative will not matter, i.e. the equations are insensitive to integration constants.

The key difference between this proposition and the previous results is that the operator $f(\partial_y)$ appearing here is not a power series operator. Instead, it is an operator that acts formally to implement translations, integrations, and differentiations. This gives rigorous justification to how equations (5) - (7) are used in practice (see [6, 7]).

3.3. Fourier-type methods

Taking the limit of the finite interval equation (9), we have the following proposition:

Proposition 3.6. *If $f : \mathbb{R} \rightarrow \mathbb{R}$ is an entire and Fourier transformable function, then*

$$\int_{-\infty}^{\infty} f(x) e^{ixy} dx = \lim_{a \rightarrow \infty} f(-i\partial_y) 2a \operatorname{sinc}(ay). \quad (35)$$

Notice that the function $\frac{a}{\pi} \operatorname{sinc}(ax)$ converges to the Dirac delta in the weak limit as $a \rightarrow \infty$, so writing equation (35) as

$$\int_{-\infty}^{\infty} f(x) e^{ixy} dx = \lim_{a \rightarrow \infty} 2\pi f(-i\partial_y) \frac{a}{\pi} \operatorname{sinc}(ay), \quad (36)$$

we can see that equation (35) is a regularized representation of equation (13).

However, as we shall later see, (13) holds in a more general sense. We start by noticing that due to the presence of the Dirac delta, equation (13) must be treated inherently as a distributional identity. Whereas equations (9) - (12) can be treated without passing through distribution theory, equation (13) cannot even be stated without the presence of a distribution. While this yields slightly more complexity, it also means that equation (13) can be generalized to hold in a much broader sense than the others. We now clarify the precise setting for equation (13) by introducing a few relevant definitions. The full technical treatment of equation (13) is presented in section 5.

Recall that a distribution is a continuous linear functional on the space of smooth, compactly supported test functions on \mathbb{R} . We denote the space of test functions by \mathcal{D} , and its continuous dual space, i.e., the space of distributions, by \mathcal{D}' . The action of a distribution $f \in \mathcal{D}'$ on a test function $\varphi \in \mathcal{D}$ is denoted by the pairing $\langle f, \varphi \rangle \in \mathbb{C}$, and a sequence of distributions $f_n \in \mathcal{D}'$ is said to converge to $f \in \mathcal{D}'$ if and only if we have

$$\lim_{n \rightarrow \infty} \langle f_n, \varphi \rangle = \langle f, \varphi \rangle \quad (37)$$

for all test functions $\varphi \in \mathcal{D}$.

One might expect that equation (13) holds as an ordinary distribution, but it is not so simple. To ensure that Fourier transforms and power series are well-defined, we must introduce an alternate space of test functions.

Let $\mathcal{Z} \equiv \mathcal{F}(\mathcal{D})$ denote the space of functions obtained from a smooth, compactly supported test function via the Fourier transform. We will say that a sequence of functions $F_n \in \mathcal{Z}$ converges to $F \in \mathcal{Z}$ if and only if the respective Fourier transforms $f_n \in \mathcal{D}$ converge to a corresponding $f \in \mathcal{D}$. We will call \mathcal{Z} the space of Paley-Wiener (PW) test functions. The continuous dual space \mathcal{Z}' will then be called the space of Paley-Wiener (PW) distributions. The spaces \mathcal{Z} and \mathcal{Z}' turn out to be the proper setting for equation (13).

In what follows, let $L^1(\mathbb{R}) + L^2(\mathbb{R})$ denote the space of functions which can be written as $f_1 + f_2$, where $f_1 \in L^1(\mathbb{R})$ and $f_2 \in L^2(\mathbb{R})$. The next proposition gives the precise meaning of equation (13) as a Paley-Wiener distribution.

Proposition 3.7. *Let $f \in L^1(\mathbb{R}) + L^2(\mathbb{R})$ be an entire function with power series expansion*

$$f(x) = \sum_{k=0}^{\infty} a_k x^k. \quad (38)$$

Then the equation

$$\hat{f}(y) = \sqrt{2\pi} f(-i\partial_y) \delta(y) \equiv \sqrt{2\pi} \sum_{k=0}^{\infty} a_k (-i\partial_y)^k \delta(y) \quad (39)$$

holds as an equality of Paley-Wiener distributions in the space \mathcal{Z}' . Explicitly, for any

Paley-Wiener test function $\varphi \in \mathcal{Z}$, we have

$$\int_{\mathbb{R}} \hat{f}(y) \varphi(y) dy = \langle f(-i\partial_y) \delta(y), \varphi(y) \rangle = \sum_{k=0}^{\infty} (-i)^k a_k \varphi^{(k)}(0). \quad (40)$$

For certain choices of f , we may evaluate (13) formally and lift the distributional identity to a proper equality of ordinary functions. The following theorem is the equivalent of Proposition 3.5 for equation (13).

Proposition 3.8. *Let $f \in L^1(\mathbb{R}) + L^2(\mathbb{R})$ be an entire function of the form*

$$f(x) = \sum_{n=1}^N \frac{c_n e^{ib_n x} + d_n e^{-ib_n x}}{x^{n_n}}, \quad (41)$$

for constants $c_k, d_k \in \mathbb{C}$, $b_k \in \mathbb{R}$, and $n_k \in \mathbb{N}_+$. Then we have

$$\hat{f}(y) = \sqrt{2\pi} f(-i\partial_y) \delta(y), \quad (42)$$

which holds as an equality of functions. The right-hand side is given by the formal action of $f(-i\partial_y)$ on $\delta(y)$, where $e^{b\partial_y}$ acts via translation by b , and where ∂_y^{-n} acts by formal anti-differentiation, giving us

$$\partial_y^{-n} \delta(y) = R_{n-1}(y) = \frac{y^{n-1}}{(n-1)!} \Theta(y), \quad (43)$$

where R_{n-1} is the generalized ramp function which defines an n th anti-derivative of the Dirac delta. Here, $\Theta(y)$ denotes the Heaviside step function.

As with Proposition 3.5, the operators $e^{b\partial_y}$ and ∂_y^{-n} acts formally (instead of as a power series operator) via translation and integration. Also like Proposition 3.5, the choice of representative for the anti-derivative does not matter. For simplicity, we have taken $\partial_y^{-n} \delta(y) = R_{n-1}(y)$, but the result holds equivalently with $\partial_y^{-n} \delta(y) = R_{n-1}(y) + p_{n-1}(y)$ for any polynomial $p_{n-1}(y)$ of degree less than or equal to $n-1$.

Finally, if f is integrable on the real line, we can evaluate $\hat{f}(y)$ at $y = 0$ to conclude that

$$\int_{\mathbb{R}} f(x) dx = \lim_{y \rightarrow 0} 2\pi f(-i\partial_y) \delta(y), \quad (44)$$

which is precisely equation (8), for functions of the form (41).

4. Example: Borwein Integral

There are two advantages of the new method. First, it allows one to evaluate certain integrals systematically without the “clever tricks” required to do the integrals in ordinary ways. Second, when the new methods are applicable they are usually much faster in comparison to ordinary methods. As an example, we now show how to use our

new methods to understand the otherwise strange and curious behavior of the Borwein integrals in a straightforward way.

Let us define the n th Borwein integral [8] as

$$B_n = \int_{-\infty}^{\infty} \prod_{k=1}^n \operatorname{sinc}\left(\frac{x}{2k-1}\right) dx = \int_{-\infty}^{\infty} \operatorname{sinc}(x) \operatorname{sinc}\left(\frac{x}{3}\right) \cdots \operatorname{sinc}\left(\frac{x}{2n-1}\right) dx. \quad (45)$$

These integrals, first studied by David Borwein and Jonathan Borwein, begin as

$$B_1 = \pi, \quad B_2 = \pi, \quad B_3 = \pi, \quad B_4 = \pi, \quad \dots \quad (46)$$

A naive guess would be that the sequence of Borwein integrals is in fact constant at π but, surprisingly, this is not so. The sequence begins to taper off at 8th term, with $B_7 = \pi$ and

$$B_8 = \pi - \frac{6879714958723010531}{467807924720320453655260875000} \pi \approx (1 - 1.47 \times 10^{-11}) \pi. \quad (47)$$

The strange behavior of this sequence, and its associated cousins, has attracted some attention in recent years, receiving a graphical proof [9], as well some generalizations [10].

To start with, let us apply equation (8) in the form of Proposition 3.8 to the sinc integral. We obtain

$$\int_{-\infty}^{\infty} \operatorname{sinc}(x) dx = 2\pi \lim_{y \rightarrow 0} \operatorname{sinc}(-i\partial_y) \delta(y) \quad (48)$$

$$= \pi \lim_{y \rightarrow 0} \frac{e^{\partial_y} - e^{-\partial_y}}{\partial_y} \delta(y) \quad (49)$$

$$= \pi \lim_{y \rightarrow 0} (e^{\partial_y} - e^{-\partial_y}) (\Theta(y) + C) \quad (50)$$

$$= \pi \lim_{y \rightarrow 0} [(\Theta(y+1) + C) - (\Theta(y-1) + C)] \quad (51)$$

$$= \pi (\Theta(1) - \Theta(-1)) \quad (52)$$

$$= \pi, \quad (53)$$

where in the second line we expressed the sinc function in terms of complex exponentials, in the third line we used the fact that the anti-derivative of the Dirac delta is the Heaviside Θ (with an undetermined constant C), and in the fourth line we used the fact that $e^{a\partial_y}$ acts formally as the translation operator $e^{a\partial_y} \varphi(y) = \varphi(y+a)$. Thus we see that equation (8) has integrated the sinc function over the real line, as claimed. We have explicitly kept the constant of integration here to show its eventual cancellation, as stated in Proposition 3.8. From this point forward, we will suppress any further integration constants.

Now let us proceed onwards to the next Borwein integral. For notational simplicity, let us write the translation operator as $e^{a\partial} = T_a$. With the same procedure as before,

we have

$$\int_{-\infty}^{\infty} \text{sinc}(x) \text{sinc}\left(\frac{x}{3}\right) dx = 2\pi \lim_{y \rightarrow 0} \text{sinc}(-i\partial_y) \text{sinc}\left(-i\frac{\partial_y}{3}\right) \delta(y) \quad (54)$$

$$= \frac{3\pi}{2} \lim_{y \rightarrow 0} (T_1 - T_{-1}) \left(T_{\frac{1}{3}} - T_{-\frac{1}{3}}\right) \iint dy \delta(y) \quad (55)$$

$$= \frac{3\pi}{2} \lim_{y \rightarrow 0} (T_1 - T_{-1}) \left(T_{\frac{1}{3}} - T_{-\frac{1}{3}}\right) R(y), \quad (56)$$

where the ramp function $R(y)$ is the anti-derivative of the Heaviside Θ function, i.e., $R(y) = y\Theta(y)$. Then we find

$$B_2 = \frac{3\pi}{2} \left[R\left(1 + \frac{1}{3}\right) + R\left(-1 - \frac{1}{3}\right) - R\left(1 - \frac{1}{3}\right) - R\left(-1 + \frac{1}{3}\right) \right]. \quad (57)$$

The ramp functions evaluated at positive arguments survive, and the rest go to zero. We are then left with

$$B_2 = \frac{3\pi}{2} \left[R\left(1 + \frac{1}{3}\right) - R\left(1 - \frac{1}{3}\right) \right] = \frac{3\pi}{2} \left(1 + \frac{1}{3} - 1 + \frac{1}{3}\right) = \pi, \quad (58)$$

which is exactly what we expected.

In general, with suitable notation, we can straightforwardly write down a general expression for the n th Borwein integral. Let $R_n(x)$ denote the n th polynomial ramp function as defined by equation (43). Adopting the notation of Borwein [1], let $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$ denote an n -tuple over $\gamma_i \in \{-1, 1\}$ and denote $\text{sign}(\gamma)$ as the product of all entries in γ . Finally, let

$$\beta_\gamma = \sum_{k=1}^n \frac{\gamma_k}{2k-1}. \quad (59)$$

Following the same procedure as before, we see that the n th Borwein integral will be expressed as a large sum over all possible combinations of shifts of different signs. With a bit of work, we can write down the integral in our notation as

$$B_n = \frac{(2n-1)!!\pi}{2^{n-1}} \sum_{\gamma \in \{-1, 1\}^n} \text{sign}(\gamma) R_{n-1}(\beta_\gamma), \quad (60)$$

where the sum is over all n -tuples $\gamma \in \{-1, 1\}^n$ and where the double factorial $(2n-1)!!$ denotes the product over odd numbers below $2n-1$, i.e. $(2n-1)!! = (2n-1) \times (2n-3) \times \dots \times 3 \times 1$.

In the sum above, we have $\beta_\gamma = -\beta_{-\gamma}$ so at most half the sum contributes due to the ramp function. Therefore, we can equivalently sum over the n -tuples $\gamma \in \{-1, 1\}^n$ whose leading term γ_1 is 1. We also have to introduce a sign correction for the γ s whose original leading term was $\gamma_1 = -1$. It follows we can write

$$B_n = \frac{(2n-1)!!\pi}{2^{n-1}(n-1)!} \sum_{\gamma; \gamma_1 > 0} \text{sign}(\gamma) \text{sign}(\beta_\gamma) \beta_\gamma^{n-1}, \quad (61)$$

where the sums over γ from this point on will be over $\gamma \in \{-1, 1\}^n$ with leading term $\gamma_1 = 1$.

Now we can distinguish between two cases. The first is when $\text{sign}(\beta_\gamma) > 0$ for all γ . This happens if and only if $1 > \frac{1}{3} + \dots + \frac{1}{2n-1}$. Using a trick from [8] we can perform the above sum explicitly in this case. Note that we have

$$\sum_{\gamma; \gamma_1 > 0} \text{sign}(\gamma) e^{\beta_\gamma x} = e^x \prod_{k=2}^n \left(e^{\frac{x}{2k-1}} - e^{-\frac{x}{2k-1}} \right). \quad (62)$$

Since

$$\left(e^{\frac{x}{2k-1}} - e^{-\frac{x}{2k-1}} \right) = \frac{2x}{2k-1} + \mathcal{O}(x^2), \quad (63)$$

we can equate the coefficients of x^{n-1} in the previous equation to get

$$\sum_{\gamma; \gamma_1 > 0} \text{sign}(\gamma) \frac{\beta_\gamma^{n-1}}{(n-1)!} = \prod_{k=2}^n \frac{2}{2k-1} = \frac{2^{n-1}}{(2n-1)!!}, \quad (64)$$

and therefore we have

$$\sum_{\gamma; \gamma_1 > 0} \text{sign}(\gamma) \beta_\gamma^n = \frac{2^{n-1}(n-1)!}{(2n-1)!!}. \quad (65)$$

Of course, this means that we have $B_n = \pi$ so long as $1 > \frac{1}{3} + \dots + \frac{1}{2n-1}$, which is satisfied for $n \leq 7$ but not for $n > 7$. This is precisely the reason that the sequence begins to break off.

In general, both the terms with $\text{sign}(\beta_\gamma) > 0$ and $\text{sign}(\beta_\gamma) < 0$ contribute to the sum and in this case we get

$$B_n = \frac{(2n-1)!!\pi}{2^{n-1}(n-1)!} \left(\sum_{\beta_\gamma > 0} \text{sign}(\gamma) \beta_\gamma^{n-1} - \sum_{\beta_\gamma < 0} \text{sign}(\gamma) \beta_\gamma^{n-1} \right) \quad (66)$$

$$= \frac{(2n-1)!!\pi}{2^{n-1}(n-1)!} \left(\sum_{\gamma; \gamma_1 > 0} \text{sign}(\gamma) \beta_\gamma^{n-1} - 2 \sum_{\beta_\gamma < 0} \text{sign}(\gamma) \beta_\gamma^{n-1} \right) \quad (67)$$

$$= \frac{(2n-1)!!\pi}{2^{n-1}(n-1)!} \left(\frac{2^{n-1}(n-1)!}{(2n-1)!!} - 2 \sum_{\beta_\gamma < 0} \text{sign}(\gamma) \beta_\gamma^{n-1} \right) \quad (68)$$

$$= \pi \left(1 - \frac{(2n-1)!!}{2^{n-2}(n-1)!} \sum_{\beta_\gamma < 0} \text{sign}(\gamma) \beta_\gamma^{n-1} \right), \quad (69)$$

where the latter sum over $\beta_\gamma < 0$ start contributing only when $n > 7$. Let us now look at the case $n = 8$, where the latter sum involves only a single term given by

$$\frac{(15)!!}{2^6 \times 7!} \left(1 - \frac{1}{3} - \frac{1}{5} - \frac{1}{7} - \frac{1}{9} - \frac{1}{11} - \frac{1}{13} - \frac{1}{15} \right)^7 = \frac{-6879714958723010531}{467807924720320453655260875000},$$

which means that B_8 has a deviation of the order 10^{-11} from π , as we saw previously.

These are precisely the results originally obtained by the Borweins in [8]. But whereas the Borweins obtained their results through clever trigonometric manipulations, the form of the Borwein integrals fall out very naturally through equation (8). Following the same type of arguments, we can obtain the following general result for minimal additional effort.

Theorem 4.1. [11] : If $a_1, \dots, a_m, b_1, \dots, b_n$, and c are positive real numbers such that

$$c > a_1 + \dots + a_m + b_1 + \dots + b_n, \quad (70)$$

then we have

$$I = \int_{-\infty}^{\infty} \text{sinc}(a_1 x) \cdots \text{sinc}(a_m x) \cos(b_1 x) \cdots \cos(b_n x) \text{sinc}(cx) dx = \pi. \quad (71)$$

Proof. : First, we can without loss of generality assume that $c = 1$ by dividing all our parameters by c . Expanding the sinc functions and the cosines in terms of complex exponentials, equation (8) allows us to write

$$I = \frac{\pi}{2^{n+m}} \left(\prod_{i=1}^m a_i \right)^{-1} \lim_{y \rightarrow 0} \prod_{i=1}^m (T_{a_i} - T_{-a_i}) \prod_{j=1}^n (T_{b_j} + T_{-b_j}) (T_1 - T_{-1}) R_m(y). \quad (72)$$

Now let us define γ similarly to before as an $(m+n)$ -tuple over $\{1, -1\}$. We will define $\text{sign}(\gamma)$ as the product of the first m entries. Likewise, define

$$\beta_\gamma = \sum_{i=1}^m \gamma_i a_i + \sum_{j=1}^n \gamma_{j+m} b_j. \quad (73)$$

Then we have

$$I = \frac{\pi}{2^{n+m}} \left(\prod_{i=1}^m a_i \right)^{-1} \sum_{\gamma \in \{1, -1\}^{n+m}} \text{sign}(\gamma) [R_m(\beta_\gamma + 1) - R_m(\beta_\gamma - 1)]. \quad (74)$$

By assumption, we have

$$1 > a_1 + \dots + a_m + b_1 + \dots + b_n, \quad (75)$$

so the term $R_m(\beta_\gamma - 1)$ must vanish for all γ while $R_m(\beta_\gamma + 1)$ will be non-vanishing for all γ . The analog of equation (62) that we need here is given by

$$\sum_{\gamma \in \{1, -1\}^{n+m}} \text{sign}(\gamma) e^{(1+\beta_\gamma)x} = e^x \prod_{i=1}^m \left(e^{\frac{x}{a_i}} - e^{-\frac{x}{a_i}} \right) \prod_{j=1}^n \left(e^{\frac{x}{b_j}} - e^{-\frac{x}{b_j}} \right). \quad (76)$$

Expanding and equating the coefficients of x^m , we must have

$$\sum_{\gamma \in \{1, -1\}^{n+m}} \text{sign}(\gamma) (1 + \beta_\gamma)^m = 2^{n+m} m! \prod_{i=1}^m a_i. \quad (77)$$

It follows that we have

$$I = \frac{\pi}{2^{n+m}} \left(\prod_{i=1}^m a_i \right)^{-1} \sum_{\gamma \in \{1, -1\}^{n+m}} \text{sign}(\gamma) R_m(\beta_\gamma + 1) \quad (78)$$

$$= \frac{\pi}{2^{n+m} m!} \left(\prod_{i=1}^m a_i \right)^{-1} \sum_{\gamma \in \{1, -1\}^{n+m}} \text{sign}(\gamma) (\beta_\gamma + 1)^m \quad (79)$$

$$= \frac{\pi}{2^{n+m} m!} \left(\prod_{i=1}^m a_i \right)^{-1} \left(2^{n+m} m! \prod_{i=1}^m a_i \right) = \pi, \quad (80)$$

where we've used the results of equation (27) in the third equality. This is precisely the result as desired. \square

Interestingly, the theorem above allows us to construct sequences of integrals which are constant at π for arbitrary lengths before eventually suddenly tapering off, simply by choosing c sufficiently large. If we choose a_n and b_n to be convergent as series, then we can have sequences of integrals which continue indefinitely to evaluate to π .

5. Proofs

In this section we prove the propositions listed in section 3.

Proof of Proposition 3.1. Let us denote the partial sums of f by f_N . Then we have:

$$\int_a^b f(x) e^{ixy} dx = \int_a^b \lim_{N \rightarrow \infty} f_N(x) e^{ixy} dx \quad (81)$$

$$= \lim_{N \rightarrow \infty} \int_a^b f_N(x) e^{ixy} dx \quad (82)$$

$$= \lim_{N \rightarrow \infty} \int_a^b f_N(-i\partial_y) e^{ixy} dx \quad (83)$$

$$= \lim_{N \rightarrow \infty} f_N(-i\partial_y) \int_a^b e^{ixy} dx \quad (84)$$

$$= \lim_{N \rightarrow \infty} f_N(-i\partial_y) \frac{e^{iby} - e^{iay}}{iy} \quad (85)$$

$$= f(-i\partial_y) \frac{e^{iby} - e^{iay}}{iy}. \quad (86)$$

The second line follows from the uniform convergence of a power series on a finite interval which allows us to exchange the order of integration and summation. The third line follows from the fact that e^{ixy} is an eigenfunction of the operator $-i\partial_y$, with eigenvalue x . The fourth line is the Leibniz rule for differentiating under the integral sign, which is valid since $(-i\partial_y)^n e^{ixy}$ exists and is continuous for all finite n . The remaining steps are straightforward.

The proof of equation (10) follows analogously by replacing e^{ixy} with e^{xy} . \square

Proof of Proposition 3.2. If the function f is Laplace transformable, then by definition the limit

$$\lim_{a \rightarrow \infty} \int_0^a f(x) e^{-xy} dx = \lim_{a \rightarrow \infty} f(-\partial_y) \frac{1 - e^{-ay}}{y} \quad (87)$$

holds for all y within the domain of convergence. The latter equality comes from an application of proposition 3.1, which holds here for functions f which are entire. This establishes equation (11). Equation (12) follows analogously. \square

Proof of Proposition 3.6. If f is Fourier transformable and entire, then the limit

$$\lim_{a \rightarrow \infty} \int_{-a}^a f(x) e^{ixy} dx = \lim_{a \rightarrow \infty} f(-i\partial_y) \frac{e^{ia y} - e^{-ia y}}{iy} \quad (88)$$

exists by definition, where the latter equality follows from proposition 3.1. Thus we have

$$\int_{-\infty}^{\infty} f(x) e^{ixy} dx = \lim_{a \rightarrow \infty} f(-i\partial_y) \frac{e^{ia y} - e^{-ia y}}{iy} = \lim_{a \rightarrow \infty} f(-i\partial_y) 2a \operatorname{sinc} ay, \quad (89)$$

as required. \square

Proof of Proposition 3.3. The proofs of all the equations are essentially identical to the proofs of propositions 3.2 and 3.6, but by taking the limit $y \rightarrow 0$ first. Note that equation (24) is simply the sum of equations (22) and (23). \square

Proof of Proposition 3.4. Let us consider taking the Laplace transform of f term-by-term. Then we get

$$\int_0^{\infty} f(x) e^{-xy} dx = \int_0^{\infty} \sum_{k=0}^{\infty} a_k x^k e^{-xy} dx. \quad (90)$$

Each of the partial sums $\sum_{k=0}^N a_k x^k e^{-xy}$ is dominated by the function

$$\sum_{k=0}^{\infty} |a_k| x^k e^{-xy} = \tilde{f}(x) e^{-xy}, \quad (91)$$

which by assumption is integrable on $[0, \infty)$ for $y > y_0$. Applying the dominated convergence theorem, we may interchange summation and integration to get

$$\int_0^{\infty} f(x) e^{-xy} dx = \sum_{k=0}^{\infty} a_k \int_0^{\infty} x^k e^{-xy} dx \quad (92)$$

$$= \sum_{k=0}^{\infty} a_k \frac{k!}{y^{k+1}} \quad (93)$$

$$= \sum_{k=0}^{\infty} a_k (-1)^k \frac{d^k}{dy^k} \frac{1}{y} \quad (94)$$

$$= f(-\partial_y) \frac{1}{y}, \quad (95)$$

with the latter series convergent for $y > y_0$.

Again, the proof of equation (12) follows analogously. \square

Before proving Proposition 3.5, we first prove a lemma showing that the power series operator $e^{b\partial}$ acts by translation on analytic functions.

Lemma 5.1. *Suppose function $f : \mathbb{R} \rightarrow \mathbb{R}$ is analytic on the open disk of radius R around x_0 and $0 < |b| < R$. Then*

$$e^{b\partial_x} f(x_0) = f(x_0 + b). \quad (96)$$

Proof. Around x_0 , f has Taylor series

$$f(x_0 + b) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(x_0) b^n. \quad (97)$$

On the other hand, by definition of $e^{b\partial_x}$,

$$e^{b\partial_x} f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} (b\partial_x)^n f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} b^n f^{(n)}(x). \quad (98)$$

Evaluated at x_0 , this gives the same result as (97). \square

Proof of Proposition 3.5. The idea of the proof is to show that the formal action of $f(-\partial_x)$ in equation (5) agrees with the results obtained by acting with $f(-\partial_x)$ as a power series operator in equation (22). Consider f of the form

$$f(-\partial_x) = \sum_{j=1}^N c_j e^{b_j \partial_x} (-\partial_x)^{n_j}, \quad (99)$$

with $b_i > 0$, $c_j \in \mathbb{C}$, and $n_j \in \mathbb{Z}$. Letting $g(x) = (1 - e^{-ax})/x$, we can write equation (22) as

$$\int_0^{\infty} f(x) dx = \lim_{a \rightarrow \infty} \lim_{x \rightarrow 0} \sum_{j=1}^N c_j e^{b_j \partial_x} (-\partial_x)^{n_j} g(x). \quad (100)$$

We will split the proof into two cases, corresponding to different signs of n_j .

First, consider the terms with $n_j < 0$. We know that $f(-\partial_x)$ is a power series operator which acts on $g(x)$ as

$$f(-\partial_x)g(x) = \sum_{k=0}^{\infty} a_k (-\partial)^k g(x). \quad (101)$$

Using the fact that integrating n times and then differentiating n times is the identity operator, we can write

$$g(x) = \partial_x^n \partial_x^{-n} g(x) = \partial_x^n g^{(-n)}(x), \quad (102)$$

where $g^{(-n)}$ is any n th anti-derivative of g . Clearly the choice of anti-derivative will not matter here since there will always be a factor of ∂_x^n acting on $g^{(-n)}$, so from now on we will choose an arbitrary representative. Thus, we can write

$$f(-\partial_x)g(x) = \sum_{k=0}^{\infty} a_k (-\partial_x)^k \partial_x^n g^{(-n)}(x) = [\partial_x^n f(-\partial_x)] g^{(-n)}(x). \quad (103)$$

Applying this result to the terms with $n_j < 0$, we get

$$\sum_{n_j < 0} c_j e^{b_j \partial_x} (-\partial_x)^{-|n_j|} g(x) = \sum_{n_j < 0} (-1)^{n_j} c_j e^{b_j \partial_x} g^{(-|n_j|)}(x) \quad (104)$$

$$= \sum_{n_j < 0} (-1)^{n_j} c_j g^{(-|n_j|)}(x + b_j), \quad (105)$$

where the last equality follows acting with $e^{b \partial_x}$ as the translation operator, justified by lemma 5.1 since all anti-derivatives of g are entire. We can write $g^{(-n)}$ explicitly as

$$g^{(-n)}(x) = h^{(-n)}(x) + \frac{e^{-ax}}{a^{n-1}} p_{n-2}(x) - \frac{1}{(n-1)!} x^{n-1} \text{Ei}(-ax), \quad (106)$$

where $h^{(-n)}$ is an n th anti-derivative for $1/x$, $p_{n-2}(ax)$ is some polynomial of degree at most $n-2$ (we define p_{-1} to be 0), and Ei is the exponential integral defined by

$$\text{Ei}(x) := - \int_{-x}^{\infty} \frac{e^{-y}}{y} dy. \quad (107)$$

Therefore we have

$$\begin{aligned} \sum_{n_j < 0} (-1)^{n_j} c_j g^{(-|n_j|)}(x + b_j) &= \sum_{n_j < 0} (-1)^{n_j} c_j \left(h^{(-|n_j|)}(x + b_j) + \frac{e^{-ax-ab_j}}{a^{|n_j|-1}} p_{|n_j|-2}(x + b_j) \right. \\ &\quad \left. - \frac{x^{|n_j|-1}}{(|n_j|-1)!} \text{Ei}(-ax - ab_j) \right). \end{aligned} \quad (108)$$

Taking the limits $x \rightarrow 0$ and $a \rightarrow \infty$ of the previous equation, we end up with

$$\lim_{a \rightarrow \infty} \lim_{x \rightarrow 0} \sum_{n_j < 0} (-1)^{n_j} c_j g^{(-|n_j|)}(x + b_j) \quad (109)$$

$$= \lim_{a \rightarrow \infty} \sum_{n_j < 0} (-1)^{n_j} c_j \left(h^{(-|n_j|)}(b_j) + \frac{e^{-ab_j}}{a^{|n_j|-1}} p_{|n_j|-2}(b_j) \right) \quad (110)$$

$$= \sum_{n_j < 0} (-1)^{n_j} c_j h^{(-|n_j|)}(b_j). \quad (111)$$

For $|n_j| > 1$, the term involving Ei vanishes as $x \rightarrow 0$ due to the presence of $x^{|n_j|-1}$. For $|n_j| = 1$, this term vanishes as $a \rightarrow \infty$ instead, with $\text{Ei}(-ab_j) \rightarrow 0$ as $a \rightarrow \infty$. Note that this is precisely what is given by the formal action of $f(-\partial_x)$ on $1/x$.

Likewise, for the terms involving $n_j \geq 0$, we have

$$\sum_{n_j \geq 0} c_j e^{b_j \partial_x} (-\partial_x)^{n_j} g(x) = \sum_{n_j \geq 0} (-1)^{n_j} c_j e^{b_j \partial_x} g^{(n_j)}(x) = \sum_{n_j \geq 0} (-1)^{n_j} c_j g^{(n_j)}(x + b_j). \quad (112)$$

The n th derivative of g is given by

$$g^{(n)}(x) = h^{(n)}(x) + \frac{e^{-ax} p_n(ax)}{x^{n+1}}, \quad (113)$$

where $h^{(n)}(x)$ is the n th derivative of $1/x$, and p_n is an n th degree polynomial. Therefore we have

$$\lim_{a \rightarrow \infty} \lim_{x \rightarrow 0} \sum_{n_j \geq 0} (-1)^{n_j} c_j e^{b_j \partial_x} (-\partial_x)^{n_j} g(x) \quad (114)$$

$$= \lim_{a \rightarrow \infty} \lim_{x \rightarrow 0} \sum_{n_j \geq 0} (-1)^{n_j} c_j \left(h^{(n)}(x + b_j) + \frac{e^{-ax - ab_j} p_n(ax + ab_j)}{(x + b_j)^{n+1}} \right) \quad (115)$$

$$= \lim_{a \rightarrow \infty} \sum_{n_j \geq 0} (-1)^{n_j} c_j \left(h^{(n)}(b_j) + \frac{e^{-ab_j} p_n(ab_j)}{(b_j)^{n+1}} \right) \quad (116)$$

$$= \sum_{n_j \geq 0} (-1)^{n_j} c_j h^{(n)}(b_j), \quad (117)$$

which again agrees with the formal action of $f(-\partial_x)$. Therefore equations (5) and (22) agree in all cases, and this proves the desired result. \square

To prepare for the proofs of propositions 3.7 and 3.8, let us first begin with a brief review of distribution theory. An ordinary distribution is a continuous linear functional on the space \mathcal{D} of smooth, compactly supported functions. The most familiar example of a distribution is of course the Dirac delta, which is just the evaluation functional at some point $y \in \mathbb{R}$:

$$\langle \delta(x - y), \varphi(x) \rangle = \varphi(y) \equiv \int_{\mathbb{R}} \delta(x - y) \varphi(x) dx, \quad (118)$$

where the last expression is how the Dirac delta is usually denoted, but whose meaning is solely given in terms of the former equation. All locally integrable functions f also define distributions as the kernel of an integral operator:

$$\langle f, \varphi \rangle \equiv \int_{\mathbb{R}} f(x) \varphi(x) dx. \quad (119)$$

Distributions which can be written in integral form against a locally integrable function are called regular distributions. The Dirac delta is the canonical example of a singular, i.e. non-regular, distribution.

Distributions enjoy many operations, such as differentiation and multiplication by x^n . For equation (13), an operation of relevance to us is the Fourier transform, which is defined for distributions by extending Parseval's theorem, i.e.

$$\langle \mathcal{F}[f], \varphi \rangle = \langle f, \mathcal{F}[\varphi] \rangle. \quad (120)$$

However, note that the Fourier transform as defined above is not well defined on ordinary distributions. The reason is that the Fourier transform of a compactly supported function is not itself compactly supported. To introduce a consistent notion of Fourier transform, the space of test functions must be enlarged to be an invariant subspace of the Fourier transform, and such a space of functions is typically given by the space of Schwartz functions \mathcal{S} , i.e. the space of smooth functions where the function and all its derivatives decay faster than x^n for all $n \geq 0$. The continuous dual space \mathcal{S}' is then called the space of tempered distributions, and it can be shown that \mathcal{S}' supports a well defined notion of Fourier transform given by equation (120).

It can be shown that the tempered distributional Fourier transform enjoys all the standard properties of the regular Fourier transform, such as the convolution theorem, and the conversion between multiplication by x and differentiation. For example, the distributional Fourier transform of the tempered distribution x^n is given by

$$\mathcal{F}[x^n](y) = \sqrt{2\pi}(-i\partial_y)^n \delta(y). \quad (121)$$

Thus, given an entire function $f(x) = \sum_{k=0}^{\infty} a_k x^k$, if we are permitted to take the distributional Fourier transform term-by-term, then we may conclude that

$$\hat{f}(y) = \sum_{k=0}^{\infty} a_k (-i\partial_y)^k \delta(y) \equiv \sqrt{2\pi} f(-i\partial_y) \delta(y), \quad (122)$$

from which equation (13) would follow. The convergence of distributions is robust enough that the term-by-term evaluation of the Fourier transform is permitted, provided that the resulting series actually converges in the sense of distributions. Unfortunately, it can be shown that the distributional series

$$\sum_{k=0}^{\infty} a_k (-i\partial_y)^k \delta(y) \quad (123)$$

actually diverges for every choice of f possible. Thus, contrary to what might be expected at first glance, equation (13) does not hold in the sense of distributions or even tempered distributions. The crux of the problem is that while test functions in \mathcal{D} or \mathcal{S} behave very well globally, they are not guaranteed to behave well locally. The terms $\partial_y^k \delta(y)$ probe the local structure of test functions while ignoring the global structure, and so the series becomes divergent[§].

To rectify this problem, we must consider a space of test functions which are locally well-behaved. This naturally leads us to consider test functions which are bandlimited, i.e. test functions with compactly supported Fourier transforms. This problem was

[§] To establish the divergence formally, we use a result known as Borel's theorem [12]. Borel's theorem says that given any sequence of $\{c_n\}_{n=0}^{\infty}$ of complex numbers, there exists a smooth, compactly supported function φ such that $\partial^n \varphi(0) = c_n$. Therefore, no matter which analytic function we choose for f , we will always be able to find a test function such that the resulting series is divergent. This means that the series diverges in the sense of distributions for every function f .

independently investigated by Gel'fand-Shilov [13, 14] and Ehrenpreis [15, 16], and we summarize their results here.

As we did in section 3, let \mathcal{Z} denote the Fourier transform of \mathcal{D} , i.e. we denote by \mathcal{Z} the space of functions whose Fourier transforms are smooth, compactly supported test functions. The Fourier transform of a compactly supported function can be holomorphically extended to the entire complex plane, and the class of functions which are the Fourier transforms of compactly supported functions, i.e. the space of bandlimited functions, is precisely characterized by the Paley-Wiener theorem.

Theorem 5.2 (Paley-Wiener[17]). *Let F be the Fourier transform of a function f which is compactly supported on the interval $[-A, A] \subseteq \mathbb{R}$. Then F is entire on \mathbb{C} and square integrable on \mathbb{R} . Moreover, there exists some constant $C > 0$ such that F satisfies the inequality*

$$|F(z)| \leq Ce^{A|\text{Im}(z)|}. \quad (124)$$

Conversely, any entire function F which is square integrable on \mathbb{R} and satisfies (124) is the Fourier transform of some function f which is compactly supported on $[-A, A]$.

The classical Paley-Wiener theorem assumes no smoothness conditions on the function f . To describe the space \mathcal{Z} we must extend the Paley-Weiner theorem to characterize the Fourier transforms of smooth, compactly supported functions. This was done by Gelfand and Shilov, as given in the following theorem.

Theorem 5.3 (Gel'fand - Shilov[13]). *Let $F \in \mathcal{Z}$ be the Fourier transform of a function $f \in \mathcal{D}$ which is compactly supported on $[-A, A] \subseteq \mathbb{R}$. Then F is entire, and for each $n \in \mathbb{N}$, there exists some constant $C_n > 0$ such that*

$$|z^n F(z)| \leq C_n e^{A|\text{Im}(z)|}. \quad (125)$$

Conversely, any entire function F which satisfies the family of inequalities (125) is the Fourier transform of a smooth test function $f \in \mathcal{D}$ which is compactly supported on $[-A, A]$.

In particular, the Fourier transform of a smooth compactly supported function is a Schwartz function, so that F is Schwartz on \mathbb{R} . In fact, the space \mathcal{Z} is a proper subspace of the Schwartz space \mathcal{S} which is dense in $L^p(\mathbb{R})$ for $1 \leq p < \infty$. The topology of \mathcal{Z} will be defined so as to make the Fourier transform a topological isomorphism, i.e. a sequence of functions $F_n \in \mathcal{Z}$ converges to $F \in \mathcal{Z}$ if and only if the corresponding Fourier transform $f_n \in \mathcal{D}$ converges to $f \in \mathcal{D}$. The continuous dual space to \mathcal{Z} will then be denoted \mathcal{Z}' . There does not seem to be a standard name for the spaces \mathcal{Z} and \mathcal{Z}' in the literature, so we will call \mathcal{Z} the space of Paley-Wiener (PW) test functions, and we will call \mathcal{Z}' the space of Paley-Wiener (PW) distributions.

Since the Fourier transform $\mathcal{F} : \mathcal{D} \rightarrow \mathcal{Z}$ is an isomorphism between \mathcal{D} and \mathcal{Z} , it follows that it also induces an isomorphism between the dual spaces \mathcal{D}' and \mathcal{Z}' . The Fourier transform of a distribution $u \in \mathcal{D}'$ will be a PW distribution $\hat{u} \in \mathcal{Z}'$ defined by

$$\langle \hat{u}, \varphi \rangle = \langle u, \hat{\varphi} \rangle \quad (126)$$

for $\varphi \in \mathcal{Z}$ and $\hat{\varphi} \in \mathcal{D}$. Note that since \mathcal{F} is a topological isomorphism, if we have a distributional series $\sum_{k=0}^{\infty} u_n \in \mathcal{D}'$ which converges to $u \in \mathcal{D}'$, then we may take the Fourier transform term-by-term to get $\sum_{k=0}^{\infty} \hat{u}_n$, and the result converges to $\hat{u} \in \mathcal{Z}'$. Thus we see that the space \mathcal{Z}' is the proper setting to formulate equation (13).

Proof of Proposition 3.7. Since $f \in L^1(\mathbb{R}) + L^2(\mathbb{R})$, it follows that f defines a regular distribution in \mathcal{D}' . Since f is also entire, its power series converges absolutely and uniformly on every compact subset of \mathbb{R} , and so the partial sums of f converges distributionally to f in \mathcal{D}' . Since the partial sums converge, we are permitted to take the distributional Fourier transform term-by-term to conclude that

$$\mathcal{F}[f](y) = \mathcal{F}\left[\sum_{k=0}^{\infty} a_k x^k\right] = \sum_{k=0}^{\infty} a_k \mathcal{F}[x^k](y) = \sqrt{2\pi} \sum_{k=0}^{\infty} a_k (-i\partial_y)^k \delta(y), \quad (127)$$

which holds as an equality of PW distributions in \mathcal{Z}' . In the above equation, we've used the fact that the distributional Fourier transform of x^n is given by $\sqrt{2\pi}(-i\partial_y)^k \delta(y)$. \square

In fact, there is no general need to restrict f to be in $L^1(\mathbb{R}) + L^2(\mathbb{R})$. The only difference then is that \hat{f} will in general not be a regular distribution.

Proof of Proposition 3.8. The operator $f(-i\partial_x)$ for $f(x)$ of the form (41) can be written as

$$f(-i\partial_x) = \sum_{k=1}^N \left[c_k e^{b_k \partial_x} + d_k e^{-b_k \partial_x} \right] \partial_x^{-n_k}. \quad (128)$$

First, let us calculate the right-hand side of (42) formally using $f(-i\partial_x)$ given above. The n th anti-derivative of the Dirac delta distribution is given by

$$\partial_x^{-n} \delta(x) = R_{n-1}(x) = \frac{x^{n-1}}{(n-1)!} \Theta(x). \quad (129)$$

The exponentials then act formally by translation, giving us

$$f(-i\partial_x) \delta(x) = \sum_{k=1}^N \left[c_k e^{b_k \partial_x} + d_k e^{-b_k \partial_x} \right] \partial_x^{-n_k} \delta(x) \quad (130)$$

$$= \sum_{k=1}^N \left[c_k e^{b_k \partial_x} + d_k e^{-b_k \partial_x} \right] R_{n_k}(x) \quad (131)$$

$$= \sum_{k=1}^N \left[c_k R_{n_k}(x + b_k) + d_k R_{n_k}(x - b_k) \right]. \quad (132)$$

On the other hand, by Proposition 3.7, we have distributionally

$$\hat{f}(x + y) = \sqrt{2\pi} \sum_{k=1}^N \left[c_k e^{b_k \partial_x} + d_k e^{-b_k \partial_x} \right] \partial_x^{-n_k} \delta(x + y) \quad (133)$$

$$= \sqrt{2\pi} \sum_{k=1}^N \left[c_k e^{b_k \partial_x} + d_k e^{-b_k \partial_x} \right] R_{n_k-1}(x + y), \quad (134)$$

where we have used the fact that $R_{n-1}(x)$ is an n th distributional anti-derivative of $\delta(x)$, i.e., $\partial_x^n R_{n-1}(x) = \delta(x)$. With the same reasoning given before in proposition 3.5, we may choose any representative for the anti-derivative since there will always be a term involving ∂_x^n which acts on $R_{n-1}(x)$. Namely, the identity continues to hold with $R_{n-1}(x) + p_{n-1}(x)$ for any polynomial $p_{n-1}(x)$ of degree at most $n - 1$. For simplicity, we continue by taking $p_{n-1} = 0$. Applying this identity to an arbitrary test function $\varphi \in \mathcal{Z}$, we get

$$\begin{aligned} \int_{\mathbb{R}} \hat{f}(x+y)\varphi(x) dx &= \sqrt{2\pi} \sum_{k=1}^N \int_{\mathbb{R}} R_{n_k-1}(x+y) \left[c_k e^{-b_k \partial_x} + d_k e^{b_k \partial_x} \right] \varphi(x) dx \quad (135) \\ &= \sqrt{2\pi} \sum_{k=1}^N \int_{\mathbb{R}} R_{n_k-1}(x+y) \left[c_k \varphi(x-b_k) + d_k \varphi(x+b_k) \right] dx. \end{aligned} \quad (136)$$

This holds for all PW test functions, so we may take a sequence φ_n which converges weakly to the Dirac delta to conclude that

$$\hat{f}(y) = \sqrt{2\pi} \sum_{k=1}^N [c_k R_{n_k}(y+b_k) + d_k R_{n_k}(y-b_k)], \quad (137)$$

where the above equation holds as an equality of functions. Comparing this expression with the result obtained earlier by acting with $f(-i\partial_x)$ formally, we therefore have

$$\hat{f}(y) = \sqrt{2\pi} f(-i\partial_y) \delta(y), \quad (138)$$

which is precisely equation (13). Taking the limit $y \rightarrow 0$ yields equation (8). \square

6. Conclusions and Outlook

In this paper we have stated and proven several results concerning the applicability of integration by differentiation formulas first stated in [6, 7].

First, Proposition 3.1 establishes equation (9) as the key result for integration on finite intervals. As long as the integrand f has a convergent power series covering the interval of integration $[a, b]$, the formula will apply.

Taking the limit of equation (9) to the case of the half-line or entire real line gives us Propositions 3.2, 3.3, and 3.6. These propositions hold for entire functions as long as the respective integrals are convergent.

We found that there are nontrivial issues regarding the domain of convergence for equations (11) and (12). The domain of convergence for the equations is generally smaller than the full domain of convergence for the Laplace transform. These solutions need analytic continuation. In Proposition 3.2, however, we gave regularized versions of these formulas which are guaranteed to be valid on the full domain of convergence for the Laplace transform. Further, Proposition 3.4 gives simple conditions for (11) and (12) to hold.

The same considerations hold for equations (6) - (7), which are the zero frequency cases of equations (11) and (12) used for integration. Since $y = 0$ is in general not a limit point of the domain of convergence, the domain of validity of equations (11) and (12) is nontrivial. Useful, therefore, for practical purposes is Proposition 3.5 which establishes that for entire, integrable functions of the form

$$f(x) = \sum_{j=1}^N c_j e^{-b_j x} x^{n_j}, \quad (139)$$

with $b_i > 0$, $c_j \in \mathbb{C}$, and $n_j \in \mathbb{Z}$, the equations (5) - (7) are valid with $e^{a\partial_y}$ acting formally as the translation operator by a , and ∂_y^n acting as repeated differentiation ($n > 0$) or integration ($n < 0$). Any undetermined integration constants arising will ultimately cancel and can, therefore, be neglected.

Finally, in Proposition 3.7 we found that equation (13) holds as a distributional identity defined on the space of Paley-Wiener test functions, i.e., the space of functions whose Fourier transforms are smooth, compactly supported test functions. If $f \in L^1(\mathbb{R}) + L^2(\mathbb{R})$ is an entire function of the form

$$f(x) = \sum_{n=1}^N \frac{c_k e^{ib_k x} + d_k e^{-ib_k x}}{x^{n_k}}, \quad (140)$$

for constants $c_k, d_k \in \mathbb{C}$, $b_k \in \mathbb{R}$, and $n_k \in \mathbb{N}_+$, then Proposition 3.8 says that equation (13) can be regarded to hold in the usual (non-distributional) sense. Here, the exponential and the derivative operators act formally by translation and differentiation/integration, with the n th anti-derivative of the Dirac delta being given by generalized ramp function

$$\partial_x^{-n} \delta(x) = R_{n-1}(x) = \frac{x^{n-1}}{(n-1)!} \Theta(x). \quad (141)$$

Regarding concrete applications of the integration by differentiation techniques, our example of the Borwein integrals clearly shows that the curious behavior of the Borwein sequence comes from the breaking of symmetry in the associated translations used in the evaluation of the integrals. Traditionally, most of the Borwein style integrals involving products of cosines and sines have been established through Fourier analytic methods, and in light of equation (13), it is perhaps unsurprising that the integration through differentiation technique would be so naturally suited to their evaluation.

Going forward, it should be interesting to expand the toolbox of our integration by differentiation methods by making use of Green's function techniques. Recall, for example, the way we use the integration by differentiation methods to evaluate $\int \sin(x)/x \, dx$ in two steps from (48) to (53). We first calculate $[1/(-\partial_y)]\delta(y)$, which, crucially, is the Green's function of $-\partial_y$. We can then straightforwardly apply $\sin(-\partial_y)$ on the result before multiplying by 2π and taking the limit of $y \rightarrow 0$.

More generally, for integrals of the form $\int g(x)/f(x) dx$ (or for the corresponding Fourier transforms), equation (8) yields:

$$\int_{-\infty}^{\infty} \frac{g(x)}{f(x)} dx = 2\pi \lim_{y \rightarrow 0} \frac{g(-i\partial_y)}{f(-i\partial_y)} \delta(y). \quad (142)$$

If we can solve the corresponding Green's function problem $f(-i\partial_y)G(y, z) = \delta(y - z)$, i.e.,

$$G(y, z) = \frac{1}{f(-i\partial_y)} \delta(y - z) \quad (143)$$

for $G(y, z)$, then the integral can be obtained from

$$2\pi \lim_{y \rightarrow 0} g(-i\partial_y)G(y, 0). \quad (144)$$

Namely, assume that the integral of a function $g(x)/f(x)$ is difficult to obtain directly, but that the Green's function for $f(-i\partial_y)$ can be found. As the further actions of $g(-i\partial_y)$ and the limit taking are straightforward, the result of the integral (or the Fourier transform) can then be obtained easily.

For example, consider the integral with $f(x) = x^2 + 1$ and $g(x) = \cos(x)$:

$$\int_{-\infty}^{\infty} \frac{\cos(x)}{x^2 + 1} dx \quad (145)$$

The integrand is a non-analytic function on the real line. Nevertheless, we may proceed heuristically as follows. Starting with equation (13), we have

$$\int_{-\infty}^{\infty} \frac{\cos(x)}{x^2 + 1} = 2\pi \lim_{y \rightarrow 0} \cos(-i\partial_y) (-\partial_y^2 + 1)^{-1} \delta(y) \quad (146)$$

We recognize the term $(-\partial_y^2 + 1)^{-1} \delta(y)$ as the Green's function for the differential operator $(-\partial_y^2 + 1)$, which is known to be

$$G(y) = \frac{1}{2} e^{-|y|}. \quad (147)$$

Notice in particular the non-analyticity at $y = 0$ which would render the power series method inapplicable. Continuing, our methods directly yield:

$$\int_{-\infty}^{\infty} \frac{\cos(x)}{x^2 + 1} = \pi \lim_{y \rightarrow 0} \cos(-i\partial_y) e^{-|y|} \quad (148)$$

$$= \frac{\pi}{2} \lim_{y \rightarrow 0} (e^{\partial_y} + e^{-\partial_y}) e^{-|y|} \quad (149)$$

$$= \frac{\pi}{2} \lim_{y \rightarrow 0} (e^{-|y+1|} + e^{-|y-1|}) \quad (150)$$

$$= \frac{\pi}{e}. \quad (151)$$

So far, all of our methods have focused on practical applications of the integration by differentiation formulas. This has motivated us to define the operator $f(r\partial_x)$ in terms of

a convergent power series. However, as we briefly mentioned at the beginning of Sec.3, the operator $f(r\partial_x)$ can be given a much broader definition using the functional calculus resulting from the spectral theorem. Namely, we may define the operator as

$$f(r\partial_x) = \mathcal{F}^{-1} M_{f(-irx)} \mathcal{F}, \quad (152)$$

where \mathcal{F} is the Fourier transform, and where $M_{f(-irx)}$ is the multiplication operator by $f(-irx)$. Note however, that the evaluation of the operator using the spectral definition requires prior knowledge of the Fourier transform and therefore also of the integral as the zero frequency limit. The spectral calculus definition is, therefore, of limited practical use for the evaluation of integrals. On the other hand, the broader validity of the spectral calculus based definition will allow us to extend the validity of the integration by differentiation formulas to a much greater class of functions, namely non-analytic functions. Such an extension would allow all of our equations to hold as distributional equalities for a large class of admissible functions. This may prove useful in the theoretical studies of quantum systems, where such distributional methods are widely applied.

Finally, let us recall that these integration by differentiation methods were originally motivated by the challenges of the Feynman path integral. A suitable generalization of our present results may make the path integral better defined and/or more straightforward to evaluate in practice. Quantum field theory is plagued by ultraviolet divergences and rigorous mathematical approaches to quantum field theory commonly define quantum fields as operator-valued tempered distributions. However, as we saw earlier, the space of Paley-Wiener distributions is a space with a naturally built in ultraviolet cutoff. It would be interesting to see if the space of Paley-Wiener distributions has a role to play in the mathematical foundations of quantum field theory.

Acknowledgements: AK acknowledges support through the Discovery Program of the National Science and Engineering Research Council (NSERC) of Canada.

References

- [1] Steven Weinberg. *The quantum theory of fields*, volume 2. Cambridge University Press, 1996.
- [2] Robin Ticciati. *Quantum field theory for mathematicians*, volume 72. Cambridge University Press, 1999.
- [3] Kazuo Fujikawa. Path-integral measure for gauge-invariant fermion theories. *Physical Review Letters*, 42(18):1195, 1979.
- [4] Edward B. Davies. *Heat kernels and spectral theory*, volume 92. Cambridge University Press, 1990.
- [5] Achim Kempf. Information-theoretic natural ultraviolet cutoff for spacetime. *Physical review letters*, 103(23):231301, 2009.

- [6] Achim Kempf, David M. Jackson, and Alejandro H. Morales. New dirac delta function based methods with applications to perturbative expansions in quantum field theory. *Journal of Physics A: Mathematical and Theoretical*, 47(41):415204, 2014.
- [7] Achim Kempf, David M. Jackson, and Alejandro H. Morales. How to (path-) integrate by differentiating. In *Journal of Physics: Conference Series*, volume 626, page 012015, 2015.
- [8] David Borwein and Jonathan M. Borwein. Some remarkable properties of sinc and related integrals. *The Ramanujan Journal*, 5(1):73–89, 2001.
- [9] Hanspeter Schmid. Two curious integrals and a graphic proof. *Elemente der Mathematik*, 69(1):11–17, 2014.
- [10] Gert Almkvist and Jan Gustavsson. More remarkable sinc integrals and sums. *arXiv:1405.1265*, 2014.
- [11] Nick Lord. 91.40 an amusing sequence of trigonometrical integrals. *The Mathematical Gazette*, 91(521):281–285, 2007.
- [12] Raghavan Narasimhan. *Analysis on Real and Complex Manifolds*. North-Holland Publishing Company, 1973.
- [13] Izrail M. Gel’fand and Georgi E. Shilov. *Generalized Functions*, volume 1. Academic Press, New York, 1964.
- [14] Izrail M. Gel’fand and Georgi E. Shilov. *Generalized Functions*, volume 2. Academic Press, New York, 1968.
- [15] Leon Ehrenpreis. Analytic functions and the fourier transform of distributions, i. *Ann. of Math*, 63:129–159, 1956.
- [16] Leon Ehrenpreis. Analytic functions and the fourier transform of distributions, ii. *Trans. Amer. Math Soc.*, 89:450–483, 1958.
- [17] Walter Rudin. *Real and complex analysis*. McGraw-Hill, New York, 1970.